

國立臺灣師範大學教育心理與輔導學系
教育心理學報，2011，42 卷，3 期，379-400 頁

發展二階段診斷工具探討學生 之統計迷思概念：以「相關」為例*

劉子鍵

國立中央大學
學習與教學研究所
師資培育中心

林怡均

國立中央大學
學習與教學研究所

統計素養是當代公民的必備能力，然而學生在學習統計時，常會產生迷思概念。相較於其他科學領域已經發展二階段診斷工具來探究學生的迷思概念，在統計教育領域中罕有類似的工具。據此，本研究針對基礎統計學中重要的概念，「相關」，編製二階段診斷工具，Two-tier Diagnostic Instrument about Correlation Misconceptions (TDICM)，並利用該工具對 982 名已修畢「相關」課程的高三學生之作答反應進行分析，以探討其「相關」理解情形，「相關」迷思概念的類型、普遍性與可能成因。研究結果顯示：(1) TDICM 的試題難度分配平均，內部一致性信度與鑑別度均佳；(2) 受測學生的相關概念理解情形不佳，部分受測學生雖然能答對 TDICM 的第一階段試題，卻因為缺乏對相關概念的完全理解，而無法答對 TDICM 的第二階段試題；(3) 根據受測學生的兩階段答案組型，顯示 TDICM 所診斷的 14 個相關迷思概念皆被學生所普遍擁有；(4) 上述 14 個相關迷思概念中有 7 個是過去研究曾探討過，有 7 個是本研究於診斷工具發展過程中發現學生常具備的迷思概念。此外，本研究亦在研究結果與討論中探討這 14 個迷思概念的可能成因與教育意涵。最後，本研究針對「二階段診斷工具」與「相關」迷思概念在統計教育研究之應用提出建議。

關鍵詞：二階段診斷工具、迷思概念、相關、統計教育

因應資訊的快速成長與擴散，具備能應用統計概念理解統計資訊（例如：統計圖表）的統計素養，已被視為是現代公民的核心能力（Garfield & Ben-Zvi, 2007; von Roten, 2006）。充分瞭解統

* 1.本論文之通訊作者為劉子鍵教授，通訊方式：320 桃園縣中壢市中大路 300 號 學習與教學研究所。

2.本研究承蒙國科會經費之補助（計畫編號：NSC94-2520-S-008-002-、NSC96-2520-S-008-007-MY2 與 NSC98-2628-S-008-001-MY3），特此致謝。

3.本研究論文之 Pilot study 與精簡版曾發表於台灣心理學會第 45 屆年會(1 頁)與 The 6th Annual Hawaii International Conference of Education (5 頁)。該兩篇論文之通訊作者與本篇研究論文之通訊作者為同一人（劉子鍵：ltc@cc.ncu.edu.tw）

計概念是有效運用統計概念與技能的先決條件 (Garfield & Ben-Zvi)。然而，學生學習統計時所面臨的困難 (例如：統計學名詞與日常使用之名詞易混淆、學習者容易以直觀的方式來詮釋統計概念與圖表等) 或統計教育工作者習慣的教學取向 (例如：強調計算、公式的推導以及專有名詞定義的背誦等)。使得許多學生常無法正確理解與應用統計概念，甚至因而產生迷思概念 (misconceptions) (Cohen, Smith, Chechile, Burns, & Tasi, 1996; Konold, 1995; Liu, Lin, & Tsai, 2009)。

學者對「迷思概念」有著不同的描述，有人用「兒童科學」(children's science) 或「先入概念」(preconceptions) 等名詞來描述「迷思概念」，也有人將其譯為「錯誤概念」(erroneous ideas) 或「另有概念」(alternative conceptions)。在統計教育領域中，「迷思概念」最常被用來表示學習者所持有與專家相異之統計概念 (Liu et al., 2009)。它意指學習者用來解釋特定事件或情境時採用之相互連結的概念，然而該組概念卻與當前廣被接受之概念相衝突 (Cohen et al., 1996; Gilbert, Osborne, & Fensham, 1982)。過去的一些研究顯示，迷思概念具有普遍擁有、不易被偵測與發覺、不易被改變、以及強烈影響後續學習等特性 (Gilbert et al., Nicoll, 2001)，因此常常是學習者學習與理解概念時最大的阻礙。

要有效促進學理解統計概念，研究者和教師必須先瞭解學生們可能具有的迷思概念 (Liu et al., 2009)。然而，評估學生的迷思概念對統計教育工作者而言是一件非常困難的事 (Hirsch & O'Donnell, 2001)。課堂中常使用的傳統選擇題測驗雖然施測與評分簡便，但其僅能測得學生對於問題的答案，而無法顯現學生回答該答案的理由 (Odom & Barrow, 1995; Peterson, Treagust, & Garnett, 1986)，因此較難有效診斷學生的迷思概念。訪談法 (Osborne & Gilbert, 1980) 與概念構圖法 (Novak & Gowin, 1984) 是研究者常用來探究學生迷思概念的方法。但這些方法的實施常需要耗費大量的時間與人力 (Liu et al., 2009; Odom & Barrow; Peterson et al., 1986; Treagust, 1986)。因此，這兩個方法並不適合教師在實務教學中使用，也不適合大樣本的調查研究 (Lee, 2007; Treagust)。

二階段 (two-tier) 診斷工具是一個同時適用於課堂教學應用與大樣本研究 (Tsai & Chou, 2002; Tsai, Chen, Chou, & Lain, 2007; Voska & Heikkinen, 2000) 且能夠有效診斷迷思概念的評量工具 (Chang et al., 2007; Lin, 2004; Tan, Goh, Chia, & Treagust, 2002; Treagust, 1986; 1988; Tsai et al., 2007)。二階段診斷工具試題的第一階段是由數個選項所組成的問題。試題的第二階段則是由數個針對第一階段之答案的可能理由所組成，這些理由包含 1 個正確理由與數個錯誤的理由，兩階段的題目皆為單選題。二階段診斷工具的第一階段問題在評量學生對於「現象的敘述性知識」的理解，第二階段問題則是在探查學生的「解釋性知識或是心智模式」(Tsai & Chou, 2002)。由於，編製者在編製此工具時是依據學生可能的迷思概念設計出第一階段的答案選項與第二階段的理由選項 (Peterson et al., 1986; Treagust; Treagust & Chandrasegaran, 2007)，因此藉由受測學生在二階段診斷工具上的特定答案組型 (所選擇之第一階段的答案與第二階段的理由之組成) 可診斷出該學生可能具有的迷思概念。

自從 Treagust (1988) 具體擬出二階段診斷工具的發展程序與步驟之後，該項工具就已經被廣泛地應用於調查各領域之學生的迷思概念 (Treagust & Chandrasegaran, 2007)。例如：化學 (Tan et al., 2002)，生物 (Lin, 2004) 與物理 (Chang et al., 2007)。目前，二階段診斷工具已被美國化學學會 (American Chemical Society) 列為概念式問題設計的推薦範例 (Tan et al.)，亦已有學術期刊 (International Journal of Science Education) 以二階段診斷工具作為專題 (Chiu, Guo, & Treagust, 2007)。然而，在統計教育領域中，鮮有診斷特定統計迷思概念的二階段診斷工具。因此，針對重

要的統計迷思概念發展出二階段診斷工具並瞭解該工具在統計教育的可用性，將會對統計教育之評量方式有重大的貢獻。

在眾多統計概念中，「相關」(correlation)被視為是基礎且重要的統計概念。「相關」不僅是基礎統計教科書或課程必定出現的概念，也是學生在學習許多進階的統計概念與方法(例如：迴歸)之前必須具有的先備概念，也因此有些國家在高中教育階段就將「相關」概念納入數學課程中。例如，台灣教育部(1995)所公布的高級中學數學課程標準即將「相關係數」單元納入高三數學的機率與統計(II)主題中，其中涵蓋的內容，包括：相關的意義，相關散佈圖的解釋、相關係數的計算，數據集中的直線趨勢等。而在目前較常被使用的高中數學教科書，例如：三民版(楊維哲、蔡聰明、吳隆盛，2001)、南一版(林福來、李恭晴、徐正梅、陳冒海、陳順宇，2003)、康熙版(李虎雄等人，2002)以及龍騰版(余文卿、李白飛，2002)之「相關係數」單元中，皆包含了相關的意義(含正相關、負相關與零相關)、散佈圖的解釋、相關係數的計算與相關強度的判斷等重要相關概念。

然而，一些探究與「相關」概念有關之迷思概念的研究卻發現，即便修完與「相關」概念有關的課程，學習者仍可能無法改變既有迷思概念或產生新的迷思概念(e.g., Liu & Garfield, 2002; Liu et al., 2009; Morris, 2001)。歸納而言，過去研究顯示學生們可能具備下列 7 個與「相關」概念有關之迷思概念：

- 相關意指因果 (Liu & Garfield, 2002; Liu et al., 2009; Morris, 1997)
- 正相關大於負相關 (Liu et al., 2009; Morris, 1997; Morris, 2001)
- 負相關等於無關 (Morris, 1997; 2001)
- 負相關的數值愈大，其關聯強度愈小 (Liu et al., 2009)
- 完全正(負)相關的圖形，一定為斜率為 1 的斜直線 (Liu et al., 2009)
- 相關係數等於斜率 (Liu et al., 2009)
- 零相關意指兩變數間毫無關係 (Liu et al., 2009)
- 完全正相關與完全負相關才是屬於直線關係(只有樣本點排列成一條斜直線時，兩變數(x, y)間才具有直線關係 (Liu et al., 2009))

本研究奠基於先前的研究結果，進一步發展「二階段相關迷思概念診斷工具」(Two-tier Diagnostic Instrument about Correlation Misconceptions, 簡稱 TDICM)並針對學過相關概念之高三生進行大規模調查研究，以瞭解國內高中生所具備與相關概念有關之統計素養。本研究的研究動機有四。首先，鑑於二階段診斷工具應用於教育領域之實用性(便於施測與詮釋)與相關概念的重要性，發展針對相關概念之二階段診斷工具有其必要性。此外，發展 TDICM 的過程中，會先利用概念構圖、訪談或開放式紙筆測驗等方法廣泛蒐集學生可能的迷思概念，此一過程可進一步發掘其他與相關有關的迷思概念。再者，透過 TDICM 的大規模施測，將有助於驗證學生是否普遍擁有各個與相關有關的迷思概念。最後，瞭解初學者擁有相關迷思概念的類別與普遍性，將有助於研究者與教學者改進統計教學 (Liu et al., 2009)。在台灣，「高三」是學生在正式課程中第一次接觸「相關」概念的階段，因此本研究以修畢「相關」課程的高三學生作為本研究的研究對象。

依據上述研究動機，本研究有以下三項研究目的：(1) 發展用以診斷「相關」迷思概念之二階段診斷工具 (TDICM)；(2) 透過 TDICM，瞭解初學相關概念者的理解情形(答題正確率)；以及(3) 透過 TDICM，探討初學者所擁有之「相關」迷思概念的類型以及其擁有「相關」迷思概念的普遍情形。藉由上述研究成果，將可用以瞭解國內高中生所具備與相關概念有關之統計素養，並提出可供提升統計素養的建議。

研究方法

一、二階段診斷工具發展流程

TDICM 的發展程序乃是以 Treagust (1988) 所提出之三階段十步驟為基礎，並針對過去研究對二階段診斷工具之質疑 (包含：Griffard & Wandersee, 2001; Taber, 1999) 進行改進。以下詳述本研究之診斷工具發展流程：

(一) 第一階段：定義診斷工具的內容

此階段的主要任務在確認待發展之 TDICM 所涵蓋的內容。此階段包括以下三個步驟：

1. 確認重要的命題知識。

本步驟包含下列主要工作：(1) 研究者針對台灣高中常使用的四種數學教科書，三民版 (楊維哲等人, 2001)、南一版 (林福來等人, 2003)、康熙版 (李虎雄等人, 2002) 以及龍騰版 (余文卿、李白飛, 2002) 中的機率與統計 (II) 之「相關係數」單元進行分析，整理出重要的「相關」概念。包含：正相關、負相關與零相關的意涵、散佈圖的解釋、相關強度的判斷與相關係數的計算。(2) 研究者進一步訪談 7 位具有十年以上高中數學科任教年資的教師，整理出教師們普遍認為重要且上課大都會提及的「相關」概念。

2. 發展概念圖。

根據前一步驟的成果，研究者以 Novak 與 Gowin (1984) 的方法發展出包含 29 項概念詞的「相關」概念圖。

3. 確認命題陳述與概念圖的內容效度。

本研究請三位高中數學教師 (教學專家) 與三位教授統計之大學教師 (學科專家) 檢視本研究前一步驟所發展出之概念圖的正確性，並依專家的意見修正概念圖，此概念圖主要將作為診斷工具發展時的參考。

(二) 第二階段：蒐集與整理學生的「相關」迷思概念

此階段的主要任務，乃透過文獻探討、專家焦點訪談法、與概念構圖法蒐集學生可能具有的「相關」迷思概念，並將所蒐集到的迷思概念整理成一份檢核表。此階段包括以下四個步驟。

1. 文獻探討與內容分析。

首先，研究者回顧 Morris (1997) 以及 Liu 與 Garfield (2002) 以及 Liu 等人 (2009) 和「相關」迷思概念有關的研究，並歸納出數項學生可能擁有之「相關」迷思概念。其次，研究者自常用之高中數學教科書，三民版 (楊維哲等人, 2001)、南一版 (林福來等人, 2003)、康熙版 (李虎雄等人, 2002) 以及龍騰版 (余文卿、李白飛, 2002) 中整理出統計學者強調學生在學習「相關」單元時可能誤解之處。此外，研究者亦參考重要的統計學專書 (例如：余民寧, 2005; 林清山, 2003)，整理出重要且常見的相關迷思概念，例如：相關等於因果。最後，「相關」為日常生活中常見的統計概念，但在大眾媒體上常受到曲解。因此，研究者自中文新聞資料庫中搜尋近十年來與「相關」有關之新聞，並整理出常見的迷思概念。

2. 訪談教學專家。

研究者訪談 3 位具有十年以上高中數學教學經驗的教師，蒐集高中生在課室學習「相關」概念過程中常見的迷思概念。

3. 應用訪談法與概念構圖法蒐集高中生的迷思概念。

本研究兼採訪談法與概念構圖法針對 25 名高三學生蒐集資料。兼採概念構圖法的原因包括：(1) 概念構圖法被認為非常適合用以評量(發掘)迷思概念(Liu et al., 2009; Roberts, 1999); (2) 概念構圖法讓學生自由地反應其原有的想法，因此不僅可用以確認學生的迷思概念更可以發掘新的迷思概念(Liu et al., 2009)。

4. 整理「相關」迷思概念的檢核表。

經過 4-6 步驟，本研究彙整出包含 14 個「相關」迷思概念之檢核表(有 7 項是過往文獻已發掘之迷思概念，另外，有 7 項是透過前述專家訪談與教科書分析所發現的。該 7 項迷思概念，經專家確認，其性質異於先前已發掘的迷思概念)。該檢核表是編製診斷工具的藍圖，讓研究者可以按照該檢核表設計並檢核所發展之 TDICM 的內容，以確保該診斷工具的內容效度。

(三) 第三階段：發展二階段開放式診斷工具並蒐集學生作答理由

二階段開放式診斷工具是 TDICM 的前身。本研究特別將發展二階段開放式診斷工具獨立成一個階段(包含三個步驟)，以詳細說明本研究在此階段採用哪些程序。

1. 撰寫第一階段試題。

本研究根據前述檢核表中的 14 個迷思概念發展出 14 題第一階段問題(用以評量學生對統計訊息之推理)，每一題皆針對某一個相關概念進行設計。為避免一些學者(例如：Griffard & Wandersee, 2001)所提出之二階段診斷工具具有容易提供答題者答題線索之缺點，研究者在此步驟中有下列兩項處置：(1) 參考 Wiersma 與 Jurs (1985) 所提出之編製選擇題測驗應注意事項，以及(2) 參考新聞報導與日常生活當中與「相關」有關的議題資料撰寫題幹與選項，使試題的問題情境異於教科書的內容，以降低試題所提供的答題線索。

2. 專家檢視與學生試做第一階段試題。

由 3 名教學專家與 3 名統計學專家檢視前一步驟所編製的第一階段試題能否達成下列三項目標：(1) 測量到所欲測量的迷思概念、(2) 內容正確、以及(3) 題目的敘述精確且通順。此外，由於學生對於選擇題測驗的試題陳述之覺察與詮釋並不會總是與試題設計者的想法一致(Hodson, 1993)且選擇題測驗多少皆需要閱讀理解上的技巧，因此不利於語文能力較差者(Taber, 1999)。為改善上述二階段診斷工具可能具有的潛在缺點，研究者請 10 名不同數學與語文程度的高中學生試做第一階段試題並表述其所覺察之各題試題的含義。爾後研究者針對專家與學生的意見修正試題。

3. 蒐集學生在第二階段試題所表述之理由。

研究者在發展完之第一階段試題的下方，加上詢問作答理由之開放性問題，形成「二階段開放式迷思概念診斷工具」。而後，本研究以該診斷工具對 138 名甫修畢相關單元的高三學生施測，以獲取學生選擇第一階段試題答案之理由。之後，研究者針對理由撰寫不清楚的學生進行事後訪談，以確保蒐集到之第二階段理由的有效性。

(四) 第四階段：編製二階段迷思概念診斷工具

此階段乃是 TDICM 發展歷程的最後一個階段，主要包括：理由選項彙整、專家效度分析、施測、與資料分析。此階段包括的四個步驟說明如下：

1. 理由選項彙整與編製。

此步驟主要是將第 10 步驟所蒐集的各題作答理由進行整理歸類後，作為 TDICM 的第二階段的理由選項。首先，研究者分析 138 名學生的施測資料，依據各題之第一階段的試題所被給予之

各類文字敘述理由，來判斷兩階段的組合（第一階段為單一選項，第二階段為針對該選項的文字敘述理由）是否能反應該題所欲診斷之迷思概念的意涵，並符合統計迷思概念的意義。若有，則認為該名學生具有該項迷思概念，並將該理由納入 TDICM 該題的第二階段理由選項中。此外，為降低 Griffard 與 Wandersee (2001) 所言，學生可依據邏輯推論或作答技巧來回答二階段診斷工具試題之可能性，本研究在編製第二階段之理由選項時有下列兩項原則：(1) 每一試題之第一階段的待選選項，至少在第二階段中有一至二項理由相對應。讓學生不至於因為察覺第一階段之某選項無對應之第二階段選項（反之亦然）而利用刪除策略作答。(2) 第二階段除包括理由選項之外，尚提供讓受測學生撰寫其它理由的欄位，以避免受測學生在無合適選項時「被迫」利用猜測策略選擇理由選項。爾後，研究者將各題的兩階段選項進行整合，完成 TDICM 初版問卷，並針對各題建置「TDICM 答案組型之迷思概念判斷表」(例如：表 1 即是針對第 8 題建置的迷思概念判斷表)，以作為判斷學生迷思概念的依據。

2. 專家效度分析。

研究者請 3 位高中數學教師與 3 位統計學專家根據「TDICM 答案組型之迷思概念判斷表」審查前步驟所編製之 TDICM 的各題試題，進行內容效度的檢核。以確認各題之兩階段答案組型能否用以診斷預定的相關迷思概念並提供修正意見。研究者依據專家的意見完成包含 14 題試題（每題測量一個相關迷思概念）之 TDICM 正式版問卷。

3. 正式施測。

本研究從台灣桃園縣 3 所學校便利取樣 22 個班級共 1,004 名高三學生進行施測，受測學生在 5 個月前皆已接受過「相關」課程的教學，並於 2 個多月前參與學科能力測驗。由於過去研究結果顯示學生的答題動機與作答速度會影響作答結果的有效性（例如：Griffard & Wandersee, 2001），因此本研究採取下列方法。首先，本診斷工具在指導語中明確告知受測學生下列事情，以促使學生認真作答、減少猜題：(1) 施測結果不會列入學校成績，和 (2) 將個別通知每位受測學生施測結果，以利學生瞭解自己的迷思概念及日後的大學聯考準備。此外，為了讓學生充分作答，並無答題時間的限制。平均而言，每位學生約花 28 分鐘作答。

4. 資料分析。

研究者根據二階段診斷工具的試題分析方法 (e.g., Lin, 2004; Tan et al, 2002; Treagust & Chandrasegaran, 2007)，對受測學生的答題正確率與迷思概念擁有情形以及試題的品質（包含信度、難度與鑑別度）進行分析。在答題正確率方面，學生須同時答對每一題的第一階段與第二階段問題，始能算是答題正確（滿分為 14 分）。在迷思概念擁有情形的診斷方面，依據本研究對「迷思概念」的操作型定義，若學生於二階段診斷工具的兩階段答案組型，符合研究者針對各題所制定之「TDICM 答案組型之迷思概念判斷表」的迷思概念型態，該名學生會被評為擁有某特定之迷思概念，由於本診斷工具的每一題皆是針對特定一個迷思概念所命題之，因此，每名學生最高擁有的迷思概念個數為 14 個。而每個迷思概念被擁有率，即是擁有該迷思概念的學生個數除上 982 名所有受測學生的個數所得的百分比。

以題 8 為例（請參見附錄一，題 8），該題主要在評量學生是否瞭解「相關係數(r)不因變數加(減)而改變」的概念，以及是否具有「相關係數會因為固定數的增減而改變」的迷思概念。根據表 1 所示，若學生的答案組型為 B (5)，意指該名學生具有正確概念。若學生的答案組型為 A (1) 或 A (3) 或 A (4) 或 B (2)，則顯示該名學生具有該迷思概念。而若學生的答案組型不符合檢核表中的任一組態，將被視為答題錯誤（不會被記為迷思概念）。由於本研究所發展之 TDICM 的兩階段皆為單選題，因此，學生的各題組型僅會有 1 組。在信度分析方面，研究者將學生在 14 題試題上的答題結果分成「對」與「錯」兩類，以 K-R 法分析 TDICM 的內部一致性信度。而 TDICM

試題難度，是根據受測學生在 TDICM 的每一題正確答對率（兩階段皆答對）計算而來。在 TDICM 的鑑別度分析方面，研究者先依據受測學生的正確答題總分，將總分在 67% 以上的學生視為高分組，總分在 27% 以下的學生視為低分組，爾後，將高分組於各題的答對率減去低分組於各題的答對率，以求得各題的鑑別度。

二、研究對象的背景

本研究以便利取樣的方式自桃園縣三所高級中學選取 1,004 名已學習過「相關」課程的高中三年級學生接受 TDICM 之施測，在剔除 22 名作答不完整（即兩階段診斷工具中之任一題的任一階段有缺答情形）的學生資料後，共有 982 名受測學生的資料接受分析。這些受測學生在性別與就讀類組上的分布相當平均。在性別方面，本研究的受測學生包含 454 名男性及 528 名女性。在就讀類組方面，本研究的受測學生包含 484 名社會組學生與 498 名自然組學生。歸納而言，本研究在性別與就讀類組等重要背景變數上的分佈相當平均且符合現況。

表 1 TDICM 答案組型之迷思概念判斷表（題 8）

一階	二階	代表的意義
A	①	具有「相關係數會因為固定數的增減而改變的迷思概念。
	③	具有「相關係數會因為固定數的增減而改變的迷思概念。
	④	具有「相關係數會因為固定數的增減而改變的迷思概念。
B	②	具有「相關係數會因為固定數的增減而改變的迷思概念。
	⑤	具有正確的概念。

結果與討論

一、各項試題之試題品質與答對率分析

本研究發展之 TDICM，其內部一致性信度為 0.62 (KR20)，符合 Nunally (1978) 所提出之良好選擇題測驗的信度標準 (0.5 以上)。TDICM 的滿分（兩階段皆答對）為 14 分，受測學生的得分介於 0 分至 13 分之間，平均得分為 6.5 分。TDICM 的難度與鑑別度之分析結果如表 2 所示。就 TDICM 的難度 (difficult indices) 而言，14 題試題的難度值分佈於 .28 至 .78 之間 (平均難度值為 .46)，其中有 7 題試題的難度介於 .40 至 .60 之間。顯見，對於本研究之受測學生而言，TDICM 的難度分配平均，有一半屬於難度適中的試題。就鑑別度 (discrimination) 而言，TDICM 的 14 題試題之鑑別度介於 .11 至 .72 之間 (平均鑑別度為 .48)，其中，有 11 題試題的鑑別度在 .40 以上，達到優良試題的標準 (Ebel & Frisbie, 1991)。

表 2 TDICM 之各項試題品質與答題正確率 (N = 982)

題目	試題品質		答題正確率	
	難度	鑑別度	第一階段答對率%	兩階段皆答對率%
1	0.33	0.66	34.2	32.5
2	0.42	0.58	70.5	42.1
3	0.42	0.62	51.9	41.8
4	0.46	0.57	50.8	46.3
5	0.61	0.51	80.9	61.0
6	0.54	0.67	59.3	53.9
7	0.40	0.42	42.4	40.2
8	0.63	0.53	68.3	63.0
9	0.53	0.48	94.0	52.9
10	0.41	0.19	50.2	41.4
11	0.28	0.11	34.4	27.6
12	0.33	0.47	41.5	32.6
13	0.78	0.35	81.2	78.3
14	0.36	0.60	40.0	36.4

就受測學生在 TDICM 的答對率分析而言，本研究發展之 TDICM 所針對的主題為統計領域中的「相關」概念。統計教育強調學生能夠「秉持正確概念進行推論，且能解讀統計資訊」(Garfield & Chance, 2000)。據此，TDICM 所定義之二階段診斷工具的第一階段為「統計訊息的解讀」；第二階段為「所依據之統計概念」。根據表 1 所示，受測學生在第一階段（解讀統計訊息）的答題正確率界於 34.2%至 94.0%之間，有 5 題試題的正確率在 50%以下；學生於兩階段（解讀統計訊息與所持之統計概念）的正確率介於 27.6%至 78.3%之間，有超過一半的試題（9 題），其正確率在 50%以下。此項結果顯示多數的試題對於受測學生而言是具有難度的。根據表 2 所示，982 名受測學生的第一階段正確率（解讀統計訊息）偏高，平均答對率為 57.1%。而 982 名高三生在兩階段皆對（解讀統計訊息與所持之概念）的答題正確率略低，平均為 46.4%。

根據表 2 所示，第一階段平均正確率高於兩階段的答對率約略 10.7%，此項結果顯示多數的受測學生或許可以透過答題技巧或憑著片斷的記憶正確解讀統計訊息，但卻因為對該項概念一知半解而無法通過第二階段的問題。此研究的結果與過往應用二階段診斷工具的施測結果相似，即學生所學習到的內容，往往是缺乏對命題與概念正確理解的片段知識而已 (Tan et al., 2002)。

二、迷思概念

表 3 為應用 TDICM 診斷 982 名受測學生所具有之迷思概念的分析結果。受測學生在 TDICM 所針對之 14 個與「相關」有關之迷思概念的擁有率介於 12.3%至 60.0%之間（平均擁有率為 30.7%）。在這 14 個迷思概念中有 7 個迷思概念（佔整體迷思概念個數 50%）是過去研究中少被探究者（請參見表 3 中註明星號者）。在本研究中這些迷思概念的擁有率介於 12.3%至 41.2%之間（平均擁有率為 24.2%）。另外，有 7 個迷思概念（佔整體迷思概念個數 50%）為過去研究（例如：Liu & Garfield, 2002; Liu et al., 2009; Morris, 1997; 2001）所曾探討過。在本研究中這些迷思概念的擁有率介於 16.2%至 60.0%之間（平均擁有率為 37.1%）。

Tan 等人 (2002) 認為當特定迷思概念之受測擁有率超過 10%時，表示該迷思概念確實普遍存在。根據 Tan 等人的標準，本研究編製之診斷工具所欲診斷的 14 個迷思概念皆可視為被受測學生們所普遍擁有。

表 3 受測學生相關迷思概念擁有情形 (N = 982)

迷思概念	迷思概念的百分比%	試題
相關的基本定義		
1. 正相關大於負相關 ^{abd}	28.1	1
2. 完全負相關 ($r = -1$) 意指兩個變數之間具有些微的關係 或是不具有任何關係 ^{abd}	27.5	6
3. 零相關意指兩變數間沒有任何關係 ^c	39.9	10
4. 兩變數為正相關時，意指 x 的值增加時， y 的值一定也會增加；反之， x 的值減少時， y 的值一定也會減少*	19.1	3
5. 相關代表因果 ^{abcd}	60.0	11
平均	34.9	
相關的圖示		
1. 無論形狀如何，散佈圖上的樣本點愈密集，其相關強度愈強*	21.9	2
2. 完全正相關 ($r = 1$) 的圖形必定為斜率為 1 的斜直線 ^d	16.2	4
3. 相關係數等於斜率 ^d	42.1	14
4. 樣本點排列成一條直線 (包含：水平線或垂直線) 即是完全相關*	12.3	5
5. 完全正相關與完全負相關才是屬於直線關係 (只有樣本點排列成一條斜直線時，兩變數 (x, y) 間才具有直線關係) ^d	45.8	7
6. 針對不同樣本所收集之變數資料可以進行相關分析*	13.6	13
平均	25.3	
相關係數的計算		
1. 相關係數會因變數加減固定數而改變*	27.9	8
2. 相關係數會因為變數單位的改變而改變*	41.2	9
3. 將兩變數調換相關係數會改變*	33.5	12
平均	34.2	
總平均	30.7	

註：* 過去研究鮮少探討之「相關」迷思概念

a Morris (1997) 研究中所探討之「相關」迷思概念

b Morris (2001) 研究中所探討之「相關」迷思概念

c Liu 與 Garfield (2002) 研究中所探討之「相關」迷思概念

d Liu 等人 (2009) 研究中發掘之「相關」迷思概念

本研究中擁有特定迷思概念個數之受測人數的次數分配如圖 1 所示。診斷結果顯示在本研究所探討之 14 個迷思概念中，每位受測學生所擁有之迷思概念的個數介於 0 至 11 之間（平均每位受測學生擁有 4.3 個迷思概念，標準差為 1.8）。其中，有六成以上（64.9%）的受測學生擁有至少 4 個迷思概念。上述分析結果顯示許多修畢相關課程之受測學生仍擁有多個與相關有關的迷思概念。

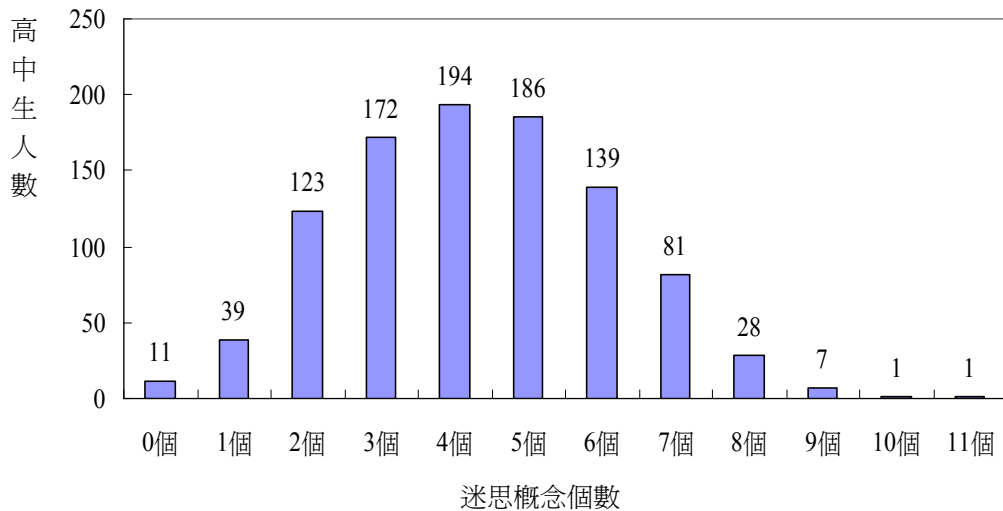


圖 1 受測學生所擁有之「相關」迷思概念個數分佈

這 14 個迷思概念分屬於相關的基本定義、相關的圖示與相關係數的計算等三個向度。受測學生在此三個向度的迷思概念平均擁有率依序為相關的基本定義（34.9%）、相關係數的計算（34.2%）以及相關的圖示（25.3%）。以下將針對此三個向度分別討論各個迷思概念的內涵與受測學生的擁有率。

（一）相關的基本定義

本研究診斷結果顯示受測學生擁有 5 個與相關的基本定義有關的迷思概念。依據擁有率之高低排序，這 5 個迷思概念分別是：（1）相關代表因果（60%）；（2）零相關意指兩變數間沒有任何關係（39.9%）；（3）正相關大於負相關（28.1%）；（4）完全負相關（ $r = -1$ ）意指兩個變數之間具有些微的關係或是不具有任何關係（27.5%）；以及（5）兩變數為正相關時，意指 x 的值增加時， y 的值一定也會增加；反之， x 的值減少時， y 的值一定也會減少（19.1%）。在 982 名受測學生中，有 885 名（91.1%）學生至少擁有一個上述與相關基本定義有關的迷思概念。

「相關代表因果」此一迷思概念是本研究所探討之各個與「相關」有關之迷思概念中受測擁有率最高者（60%受測學生擁有），也是過去文獻中最常被提及者（例如：Liu & Garfield, 2002; Liu et al., 2009; Morris, 1997）。由此可見此一迷思概念的重要性。此外，受測學生的作答反應組型顯示，學生對於相關是否代表因果之判斷會受到兩變數間相關係數之大小的影響。有 13.6% 的受測學生認為相關係數大於零即代表二變數間具有因果關係（Q11, A2）；有 44.3% 的學生是因為題目列出二個變數的相關係數（0.89）夠大，因此認為二變數間有因果關係（Q11, A1）；有 2.1% 的學生認為

題目所列的相關係數 (0.89) 不夠大, 還需要更大一些才能因此判斷二個變數間有因果關係 (Q11, B4)。(請參見附錄一, 題 11)

「零相關意指兩變數間沒有任何關係」此一迷思概念為近四成 (39.9%) 受測學生所擁有。一般而言, 教科書中常會特別以圖示的方式說明直線相關中的零相關是指二變數之間不存在「直線關係」(例如: 李虎雄等人, 2002)。然而, 本研究的診斷結果顯示一些受測學生面對兩個變數間為零相關的問題情境時, 傾向認為此二個變數間是「沒有任何關係」, 而非「沒有直線關係」。Liu 等人 (2009) 的研究提出學生之所以會有此迷思概念是因為學生常傾向援用「零」在日常生活中的意義—「沒有任何」來詮釋零相關此一統計概念。因此, 一些學生即便是學習過零相關, 仍然直觀地認為零相關代表二個變數之間沒有任何的關係。

「正相關大於負相關」與「完全負相關 ($r = -1$) 意指兩個變數之間具有些微的關係或是不具有任何關係」此二迷思概念與「相關強度」的概念有關。本研究診斷結果顯示多達 46.3% 的學生擁有此類迷思概念之一。過去, 一些小樣本的研究結果亦曾顯示學生擁有上述兩個與「相關強度」有關的迷思概念 (例如: Liu et al., 2009; Morris, 1997)。本研究利用大樣本的調查研究, 進一步證明這兩個與「相關強度」有關之迷思概念的普遍性。

在教導與學習「相關」概念時, 「相關強度」是個非常重要的概念。然而, 在本研究中, 學生的作答結果顯示許多學生並未真正理解相關係數之絕對值與「相關強度」的關係而仍然認為相關係數為正時的「相關強度」必定大於相關係數為負時的「相關強度」(例如: 有 28.1% 的受測學生認為 $r = 0.8$ 的相關強度大於 $r = -0.9$) 或認為相關係數為負時, 其值愈大代表「相關強度」愈小 (例如: 有 27.5% 的受測學生認為 $r = -1$ 代表二變數間的相關強度接近零)。Liu 等人 (2009) 曾透過訪談與概念構圖的方式發現導致學生擁有與「相關強度」有關之迷思概念的主因是學生常會引用過去已習得之數學中正數與負數的概念來不當詮釋「相關強度」的意涵。例如: 學生認為正數大於負數, 因此推論正相關大於負相關; 又如: 學生認為負數的數值愈大其值越小且 r 介於 -1 至 1 之間, 因此推論 $r = -1$ (完全負相關) 意指兩變數間具有些微的關係或是不具有任何關係。因此, 未來在統計教學中解釋「相關強度」的意義時應特別引導學生區辨「相關強度」與數學之正數負數之概念的不同。

最後, 「兩變數為正相關時, 意指 x 的值增加時, y 的值一定也會增加; 反之, x 的值減少時, y 的值一定也會減少」此一迷思概念與正相關的定義有關。過去未有研究探討此類迷思概念, 本研究的診斷結果顯示有 19.1% 的受測學生認為「正相關是指 x 的值增加 (減少) 時, y 的值一定也會增加 (減少)」並因此無法從幾組數對資料中正確地判斷哪些數對資料可能是正相關。擁有此迷思概念的學生無法正確理解當兩變數間為正相關時, 代表兩變數之數對資料間是具有「可能」關係而非「必然」關係。一般統計教科書或教師教學時, 常將正相關界定為 x 值增加 (減少) 時, y 值會有增加 (減少) 的「趨勢」或「現象」; 而將負相關界定為 x 值增加 (減少) 時, y 值會有減少 (增加) 的「趨勢」或「現象」。趨勢代表有此傾向但並不必然如此。或許是人類有忽略例外的天性 (Nisbett & Ross, 1980), 也或許「趨勢」或「現象」一詞過於抽象而不易理解, 因此讓一些學過正負相關定義之學生仍誤認為正 (負) 相關代表各數對間具有一致的對應關係。

(二) 相關的圖示

學生作答結果顯示受測學生擁有 6 個與相關的圖示有關之迷思概念。這 6 個迷思概念 (依據擁有率之高低排序) 分別是: (1) 完全正相關與完全負相關才是屬於直線關係 (只有樣本點排列成一條斜直線時, 兩變數 (x, y) 間才具有直線關係) (45.8%); (2) 相關係數等於斜率 (42.1%); (3) 無論形狀如何, 散佈圖上的樣本點愈密集, 其相關強度愈強 (21.9%); (4) 完全正相關 ($r=1$)

的圖形必定為斜率為 1 的斜直線 (16.2%); (5) 針對不同樣本所收集之變數資料可以進行相關分析 (13.6%); 以及 (6) 樣本點只要呈現一條直線 (包含: 水平線或垂直線) 即是完全相關 (12.3%)。

982 名受測學生中, 有 795 名 (81%) 學生至少擁有一個上述與相關圖形有關的迷思概念。過去, Estepa 與 Batanero (1996) 的研究結果亦顯示學生常僅根據散佈圖中部份的訊息而做出錯誤的推論。圖形常用以傳達統計訊息 (von Roten, 2006), 但學生很可能對圖形表徵具有迷思概念, 而對訊息產生錯誤解讀, 此一現象非常值得統計教育學者的重視。

在上述六個迷思概念中, 「只有樣本點排列成一條斜直線時, 兩個變數 (x, y) 之間才具有直線性關係」的被擁有率最高 (有 45.8% 的受測學生擁有此迷思概念)。雖然教科書通常會解釋線性關係代表二變數之間具有直線趨勢 (例如: 李虎雄等人, 2002), 許多受測學生似乎會過度窄化了線性關係的意涵 (認為線性關係是指二變數之間為一條直線), 並且望文生義而認為「僅有散佈圖中樣本點排列成一條斜直線時, 兩變數 (x, y) 間才具有直線關係」。

其次, 本研究發現一些受測學生會誤認相關係數與斜率之間具有某種關係, 包含「相關係數等於斜率」(42.1%) 或「完全正相關 ($r=1$) 的圖形必定為斜率為 1 的斜直線」(16.2%)。此一研究結果與 Liu 等人 (2009) 的研究結果近似。當兩變數為標準化時, 完全正相關的散佈圖中樣本點的分佈是由左下至右上呈 45 度角的斜直線。此時, 斜直線的斜率正好等於完全正相關的相關係數 ($r=1$)。學生常因此誤認完全正相關的圖形必定為斜率為 1 的斜直線, 並類推「相關係數與散佈圖之斜率間存在某種關係」。本研究的調查結果顯示, 約有 24.0% 的受測學生認為斜率愈大者代表其相關強度愈強 (Q14, A3), 而有 18.1% 的受測學生認為愈接近斜率為 1 的斜直線相關強度愈強 (Q14, B1)。(請參見附錄一, 題 14)

「無論形狀如何, 散佈圖上的樣本點愈密集, 其相關強度愈強」這個迷思概念是過去研究中較少被提及的。本研究中, 有 21.9% 的學生認為樣本點排列成緊密圓形的散佈圖, 其相關強度最強。這些學生似乎以樣本點間「疏密」的程度來判斷相關的強度, 而忽略了樣本點所排列之「形狀」的重要性。

此外, 相關涉及對同一樣本的兩個變數進行雙變數分析, 因此進行相關分析的資料必須來自同一樣本在兩變數上的數對資料。然而, 本研究中有 13.6% 的受測學生直覺地認為可藉由同一樣本之非數對資料的分佈 (例如: 同一班級之數學成績與英文成績的長條圖分佈) 判斷兩變數之相關程度。尤其當代表兩個變數之長條圖有類似的分佈時 (圖 2), 受測學生傾向認為二個變數間具有正相關。

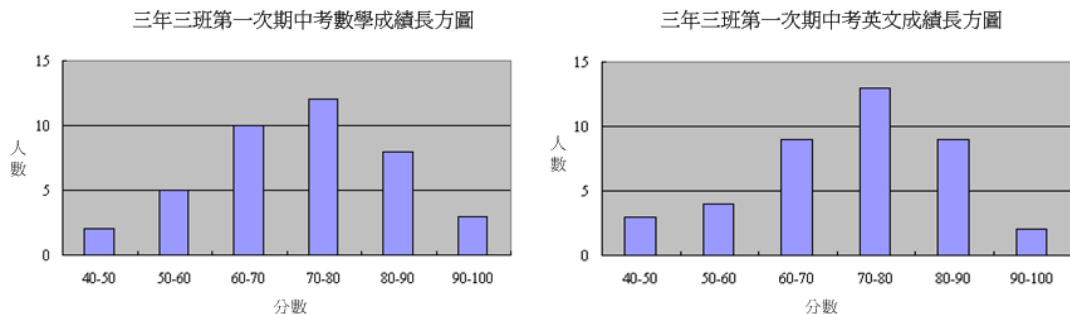


圖 2 同一樣本之非數對資料範例

最後, 約有 12.3% 的受測學生認為樣本點排列成一直線 (水平線或垂直線), 即是完全正相關。此迷思概念也是過去研究中鮮少被提及的。事實上, 完全正 (負) 相關的正確意涵是指從二變數

中的任一變數可以完全預測到另一變數。此時，散佈圖中的樣本點呈現左下至右上（完全正相關）或右上至左下（完全負相關）的斜直線。有此一迷思概念的受測學生似乎將完全正（負）相關的散佈圖為一直線的概念過度類推，而認為只要當散佈圖中樣本點呈現為一直線時即代表兩變數間呈完全相關，而忽略了當樣本點呈現垂直線或水平線時，並無法從二變數中的任一變數預測到另一變數之事實。

（三）相關係數的計算

雖然統計學專書（例如：余民寧，2005；林青山，2003）中介紹了相關係數之計算的原則和概念，然而，學生作答結果顯示，即便是修完統計相關概念的學生仍普遍擁有 3 個與相關係數計算有關的迷思概念。依據擁有率之高低排序，這 3 個迷思概念分別是：（1）相關係數會因變數單位的改變而改變（41.2%）；（2）將兩變數調換相關係數會改變（33.5%）；以及（3）相關係數會因變數加減固定數而改變（27.9%）。診斷結果顯示在 982 名受測學生中，有 671 名（68.3%）學生至少擁有一個與相關係數之計算有關的迷思概念。雖然相關係數的計算為基礎統計的教學重點，受測學生作答反應顯示學生熟練公式的計算並不代表學生對相關係數的計算擁有正確的概念。

約有三分之一（33.5%）的受測學生認為「兩變數的次序調換會改變兩變數間的相關係數」。研究者在診斷工具編製的第 6 項步驟（應用訪談法與概念構圖法蒐集高中生的迷思概念）發現此一迷思概念。當研究者談及 x 與 y 的相關與談及 y 與 x 的相關時有些學生會有不同的反應。此又尤以研究者以具體的變數名稱詢問學生時最為明顯（例如：本診斷工具的試題是以「運動時間」與「健康指數」代表 x 或 y ）。受測學生的作答反應顯示擁有此一迷思概念者可能持有的理由有三：約 19.9% 的受測學生認為兩變數對調，理應會改變相關係數（Q12, A1）。約有 2.3% 的受測學生認為相關係數為前項/後項，因此當前後項改變時相關係數自然改變（Q12, A2）。另外，有 11.3% 的受測學生認為前項為因、後項為果，當前項與後項對調時，意味著因果關係對調，因此會影響相關係數（Q12, A3）。（請參見附錄一，題 12）

其次，「相關係數會因變數加減固定數而改變」與「相關係數會因為變數單位的改變而改變」此二個迷思概念顯示受測學生並未清楚理解相關係數之定義公式的意涵。在高中數學教科書中，相關係數的計算公式會將 x 變數與 y 變數轉換成標準分數（例如：李虎雄等人，2002）。轉換成標準分數後的 x 變數與 y 變數的平均數為 0，而標準差（單位）為 1。此時，無論變數加減固定數或是變數單位的改變皆不會影響其所計算出的相關係數。但本研究的診斷結果發現，雖然受測學生所使用的教科書上皆有相關係數的公式且多數高中數學科老師在教學時亦會說明公式的意義並要求學生熟記該公式，多數學生在思考相關係數時並未從標準分數的概念著眼，而是以直觀的方式來解釋，因此導致至少有 54.8% 的受測學生擁有上述二個迷思概念之一。例如：當問及每位學生的國文成績都加五分之後，學生之國文成績與英文成績間的相關係數會產生何種變化。有 11.7% 的學生會直觀地認為增加國文成績會使相關係數變大（題 8, A1）；有 3.2% 的學生認為僅增加國文成績 5 分，不足以影響相關係數，若再增加多一些分數，兩變數間的相關係數會變大（題 8, B2）。有 3.6% 的學生認為僅加國文成績，會使英文成績相對變低，因此使相關係數由正轉負（題 8, A3）。而亦有 9.4% 學生認為僅加國文成績會使相關係數變小（題 8, A4）。（請參見附錄一，題 8）

由於，相關係數之計算常為基本統計的教學重點之一，本研究所列舉之有關相關係數計算之三迷思概念皆為過去文獻較少探討者，值得吾人重視。

結論與建議

擁有特定統計迷思概念不但不利於該概念之理解與運用，往往也會妨礙學生對進階統計概念的學習 (Liu et al., 2009)，亦會影響統計素養的提升。因此，對於實際教學者與研究者而言，瞭解學生可能擁有之統計迷思概念及其普遍程度是一件非常重要的事。然而，評估學生的統計迷思概念十分不易 (Hirsch & O'Donnell, 2001)，因為學生的迷思概念往往隱藏在其深層的知識結構之中，只有在面對特定事例時才會顯現出來。雖然，二階段診斷工具被視為是診斷迷思概念的有效工具，且已普遍被應用到不同的科學教育領域去偵測與特定重要概念有關的迷思概念。但在統計教育領域中，罕有研究發展二階段診斷工具來診斷學生的迷思概念。

「相關」是統計學中基礎且重要的概念。而過去的研究亦顯示學生即便學習過「相關」概念，仍可能擁有與「相關」有關的迷思概念 (e.g., Liu & Garfield, 2002; Liu et al., 2009)。然而，目前卻鮮有研究大規模地診斷學生可能擁有與「相關」有關的迷思概念及其普遍性。據此，本研究的主要目的乃在發展用以診斷「相關」迷思概念的二階段診斷工具，TDICM。並利用該工具進行大樣本的調查研究，瞭解初學者擁有與「相關」有關之迷思概念類型與普遍性。以下將歸納本研究之研究結論並提出對未來研究與教育應用的建議。

在二階段診斷工具之發展方面。研究者以 Treagust (1988) 所提出，用以建置二階段概念診斷工具的 3 階段 10 步驟為基礎，並參考過去研究對二階段診斷工具的質疑，例如：容易提供答題線索、不利語文程度較差者、容易受到答題者的動機與作答速度的影響 (Griffard & Wandersee, 2001; Taber, 1999) 採取相對應的試題編製措施編製 TDICM。歸納研究結果，TDICM 具有以下特性：首先，TDICM 的難度分佈平均，內部一致性信度與試題鑑別度均佳。其次，TDICM 可以診斷出學生的統計迷思概念。在本研究中，TDICM 的每一試題之第一階段的問題，皆被設計用以評量受測學生對於各種形式統計訊息 (文字、數據、圖形、符號) 的解讀能力，第二階段的問題則被設計用以評量受測學生於統計訊息解讀所奠基的統計概念。因此，受測學生在特定試題上的特定答案組型 (統計訊息解讀與所奠基的統計概念) 可顯示該名學生的迷思概念。上述設計與考量，讓 TDICM 可以有效診斷出課堂中不易察覺的相關迷思概念。

再則，作為一個統計教育的研究工具，TDICM 可用以驗證過去研究所發現的統計迷思概念並挖掘新的統計迷思概念。本研究將 Liu 等人 (2009)，Morris (1997, 2001) 所發掘之 7 個學生可能擁有之相關迷思概念納入本研究二階段診斷工具編製的考量，並經大樣本調查確認這 7 個迷思概念確實為受測學生們所普遍擁有。另外，本研究在診斷工具編製歷程中經由訪談與二階段開放式診斷工具等歷程蒐集到 7 個過去文獻鮮少提及，而學生可能擁有之迷思概念，並經大樣本調查確認這 7 個迷思概念亦確實為學生們所普遍擁有。然而，值得注意的是 TDICM 之所以具有挖掘新迷思概念的潛能，是因為該類型工具在編製的過程中採用開放性蒐集資料的方式。一旦工具編製完成，封閉式的二階段診斷工具將不復此一特性。未來研究可以持續以其他開放性蒐集資料方式探討其他可能的相關迷思概念。最後，作為一個統計迷思概念診斷工具，TDICM 具有便於評分與詮釋的特性，此點可讓該工具同時適用於大樣本的調查研究與課堂中教師即時診斷學生理解狀態。

鑑於上述 TDICM 的特性，本研究建議未來研究可參考本研究發展 TDICM 的步驟，針對其他常見之統計迷思概念 (例如：平均數、機率、中央極限定理)，發展診斷特定統計迷思概念之二階段診斷工具。

在統計教學應用方面。「相關」是公認重要的統計概念，但本研究發現初學者（高中年級學生）的「相關」概念理解情形普遍不佳。據此，研究者對未來的統計教學有下列兩項建議：（1）統計教學應該更側重統計概念的正確理解與正確運用；（2）在評量學生們的統計能力時，除了評量學生們能否正確的解讀統計訊息之外，應更深一層的評量學生是否依據正確的統計概念，以利確實掌握學生的統計能力。

此外，在「相關」迷思概念的擁有情形方面，本研究發現至少有 14 個迷思概念為修畢相關單元之受測學生所普遍擁有。這些迷思概念歸屬於相關基本概念、相關之圖示以及相關係數之計算等三類與相關有關的重要概念。雖然，這三類相關概念非常基本且是各教科書必提的統計概念，但本研究顯示在這三類概念上，有非常高比例的學生擁有迷思概念。本研究在結果與討論中曾探討這些迷思概念的內涵與可能的成因。建議未來編纂統計教材與進行統計教學時，能將這些迷思概念的內涵與可能的成因納入考量，以利避免學生建構這些迷思概念、診斷學生是否具備這些迷思概念、以及進一步設法改變學生已經具有的迷思概念以提升其與相關概念有關的統計素養。

在許多國家（例如：台灣），高中階段不僅是初學相關概念的階段也可能是必修統計概念之最後階段，若這些擁有迷思概念的學生，在未來並未升學、或升大學後並未選修統計學、或選修統計學後仍未發覺其所擁有之迷思概念，則這些學生們將極有可能繼續錯誤理解與應用相關概念，甚至奠基在此迷思概念上錯誤學習進階統計（如：迴歸）。因而如何在高中（職）階段有效診斷與改變迷思概念是重要的研究議題。本研究的作者亦以本研究的研究結果為基礎，針對「相關」概念發展「以模擬為基礎的輔助統計學習軟體」，期望未來能應用該項軟體有效改善學生的「相關」迷思概念，增進學生對於「相關」概念的理解以提升其相關概念之統計素養。

參考文獻

- 李虎雄、陳昭地、黃登源、李政貴、林祜堂、儲啓政（2002）：**高級中學數學乙（上冊）**。台北：康熙。
- 余文卿、李白飛主編（2002）：**高級中學數學甲（上冊）**。台北：龍騰。
- 余民寧（2005）：**心理與教育統計學**。台北：三民。
- 林清山（2003）：**心理與教育統計學**。台北：東華。
- 林福來、李恭晴、徐正梅、陳冒海、陳順宇（2003）：**高級中學數學乙（上冊）**。台北：南一。
- 教育部（1995）：**高級中學數學課程標準**。取自教育部網站：
http://www.find.org.tw/0105/news/0105_news_disp.asp?news_id=2735，2009年6月30日。
- 楊維哲、蔡聰明、吳隆盛（2001）：**高級中學數學甲（上冊）**。台北：三民。
- Chang, H. P., Chen, J. Y., Guo, C. J., Chen, C. C., Chang, C. Y., Lin, S. H., & Su, W. J. (2007). Investigating primary and secondary students' Learning of physics concepts in Taiwan. *International Journal of Science Education*, 29(4), 465-482.

- Chiu, M. H., Guo, C. J., & Treagust, D. F. (2007). Assessing students' conceptual understanding in science: An introduction about a national project in Taiwan. *International Journal of Science Education, 29*(4), 379-390.
- Cohen, S., Smith, G. A., Chechile, R., Burns, G., & Tasi, F. (1996). Identifying impediments to learning probability and statistics from an assessment of instructional software. *Journal Educational and Behavioral Statistics, 21*(1), 35-54.
- Ebel, R. L., & Frisbie, D. A. (1991). *Essentials of educational measurement* (5th ed.). Englewood Cliffs, NJ: Prentice-Hall.
- Estepa, A., & Batanero, C. (1996). Judgments of correlation in scatter plots: Students' intuitive strategies and preconceptions. *Hiroshima Journal of Mathematics Education, 4*, 25-41.
- Garfield, J., & Chance, B. (2000). Assessment in statistics education: Issues and challenges. *Mathematics Thinking and Learning, 2*(1&2), 99-125.
- Garfield, J., & Ben-Zvi, D. (2007). How students learn statistics revisited: A current review of research on teaching and learning statistics. *International Statistical Review, 75*(3), 372-396.
- Gilbert, J. K., Osborne, R. J., & Fensham, P. J. (1982). Children's science and its consequences for teaching. *Science Education, 66*(4), 623-633.
- Griffard, P. B., & Wandersee, J. H. (2001). The two-tier instrument on photosynthesis: What does it diagnose? *International Journal of Science Education, 23*(10), 1039-1052.
- Hirsch, L. S., & O'Donnell, A. M. (2001). Representativeness in statistical reasoning: Identifying and assessing misconceptions. *Journal of Statistics Education, 9*(2). Retrieved August 18, 2008, from <http://www.amstat.org/publications/jse/v9n2/hirsch.html>.
- Hodson, D. (1993). In search of a rationale for multicultural science education. *Science Education, 77*(6), 685-711.
- Konold, C. (1995). Issues in assessing conceptual understanding in probability and statistics. *Journal of Statistics Education, 3*(1). Retrieved August 18, 2008, from <http://www.amstat.org/publications/jse/v3n1/konold.html>.
- Lee, S. J. (2007). Exploring students' understanding concerning batteries? Theories and practices. *International Journal of Science Education, 29*(4), 497-516.
- Lin, S. W. (2004). Development and application of a two-tier diagnostic test for high school students' understanding of flowering plant growth and development. *International Journal of Science and Mathematics Education, 2*(2), 175-199.
- Liu, H. J., & Garfield, J. B. (2002). Sex differences in science statistical reasoning. *Bulletin of Educational Psychology, 34*(1), 123-138.

- Liu, T. C., Lin, Y. C., & Tsai, C. C. (2009). Identifying misconceptions about statistical correlation and their possible causes among high school students: An exploratory study using concept mapping with interviews. *International Journal of Science and Mathematics Education*, 7(4), 791-820.
- Morris, E. J. (1997). *An investigation of students' conceptions and procedural skills in the statistical topic correlation* (CITE Report. No. 230). UK: Centre for Information Technology in Education. Institute of Educational Technology. The Open University.
- Morris, E. (2001). The design and evaluation of link: A computer-based learning system for correlation. *British Journal of Educational Technology*, 32(1), 39-52.
- Nicoll, G. (2001). A report of undergraduates' bonding misconceptions. *International Journal of Science Education*, 23(7), 707-730.
- Nisbett, R., & Ross, L. (1980). *Human inference: Strategies and short-comings of social judgment*. Englewood Cliffs, NJ: Prentice-Hall.
- Novak, J. D., & Gowin, D. B. (1984). *Learning how to learn*. New York: Cambridge University Press.
- Nunally, J. C. (1978). *Psychometric Theory*. New York: McGraw-Hill.
- Odom, A. L., & Barrow, L. H. (1995). Development and application of a two-tier diagnostic test measuring college biology students' understanding of diffusion and osmosis after a course of instruction. *Journal of Research in Science Teaching*, 32(1), 45-61.
- Osborne, R. J., & Gilbert, J. K. (1980). A technique for exploring students' view of the world. *The Institute of Physics*, 15(6), 376-379.
- Peterson, R., Treagust, D., & Garnett, P. (1986). Identification of secondary students' misconceptions of covalent bonding and structure concepts using a diagnostic instrument. *Research in Science Education*, 16(1), 40-48.
- Roberts, L. (1999). Using concept maps to measure statistical understanding. *International Journal of Mathematical Education in Science and Technology*, 30(5), 707-717.
- Taber, K. S. (1999). Ideas about ionisation energy: A diagnostic instrument. *School Science Review*, 81(295), 97-104.
- Tan, K. C. G., Goh, N. K., Chia, L. S., & Treagust, D. F. (2002). Development and application of a two-tier multiple choice diagnostic instrument to assess high school students' understanding of inorganic chemistry qualitative analysis. *Journal of Research in Science Teaching*, 39(4), 283-301.
- Treagust, D. F. (1986). Evaluating students' misconception by means of diagnostic multiple choice items. *Research in Science Education*, 16(1), 199-207.
- Treagust, D. F. (1988). Development and use of diagnostic tests to evaluate students' misconception in science. *International Journal of Science Education*, 10(2), 159-169.

- Treagust, D. F., & Chandrasegaran, A. L. (2007). The Taiwan national science concept learning study in an international perspective. *International Journal of Science Education, 29*(4), 391-403.
- Tsai, C. H., Chen, H. Y., Chou, C. Y., & Lain, K. D. (2007). Current as the key concept of Taiwanese students' understandings of electric circuits. *International Journal of Science Education, 29*(4), 483-496.
- Tsai, C. C., & Chou, C. (2002). Diagnosing students' alternative conceptions in science. *Journal of Computer Assisted Learning, 18*(2), 157-165.
- von Roten, F. C. (2006). Do we need a public understanding of statistics? *Public Understanding of Science, 15*(2), 243-249.
- Voska, K. W., & Heikkinen, H. W. (2000). Identification and analysis of student conceptions used to solve chemical equilibrium problems. *Journal of Research in Science Teaching, 37*(2), 160-176.
- Wiersma, W., & Jurs, S. G. (1985). *Educational measurement and testing*. Newton, MA: Allyn & Bacon.

收稿日期：2009年03月04日

一稿修訂日期：2009年07月13日

二稿修訂日期：2009年08月05日

接受刊登日期：2009年08月05日

附錄一

二階段相關迷思概念診斷工具例題

以下呈現 4 題本文所提及之二階段相關迷思概念診斷工具的試題，每一題皆由兩階段組成，每一階段的正確答案以*標註。

題 8：

- () 某一班學生的國文成績與英文成績的相關係數(r)為 0.4。因為該班的國文成績不甚理想，因此國文老師決定將全班學生的國文成績都加上 5 分。請問，調整之後的國文成績與未調整之英文成績間的相關係數可能如何？
- (A) $r \neq 0.4$
 (B) $r = 0.4$ *
- () 為什麼我會選擇上述的答案呢？下列哪一個選項最符合我的想法：
- ① 增加每位學生的變項分數（例如：國文成績），會使兩變項間的相關係數(r)變大
 ② 因為每位學生的變項分數（例如：國文成績）的增加分數(5分)太少，故不會影響兩變項間的相關係數(r)。若增加的分數多些，則兩變項間的相關係數(r)會變大
 ③ 僅國文成績有加分，英文成績便相對降低了，因此兩變項的關係會變成負相關
 ④ 國文成績提高 5 分相對於英文成績降低 5 分，因此兩項分數間的相關係數(r)會變小
 ⑤ 增加每位學生的變項分數（例如：國文成績），並不會影響兩變項間的相關係數(r)*
 ⑥ 其它_____

題 11：

- () 一份調查研究發現：學生的出席率與學業成績間的相關係數(r)為 0.89。你是否因此斷言：學生的出席率是導致其學業成績好壞的重要原因。
- (A) 贊成
 (B) 反對*
- () 為什麼我會選擇上述的答案呢？下列哪一個選項最符合我的想法：
- ① $|r| = 0.89$ ，意指出席率與學業成績之間具有高度相關，因此有高度因果關係
 ② 相關係數(r)大於 0，意指出席率與學業成績之間具有相關，因此有因果關係
 ③ 相關表示出席率與學業成績間有直線關係，但不一定有因果關係*
 ④ 相關係數(r)要比 0.89 更大一些，才能說二者有因果關係
 ⑤ 其它_____

題 12 :

() 某運動科學研究室宣稱：運動時間與健康指數間的相關係數(r)為 0.5 如果該研究結果正確，請問健康指數與運動時間的相關係數(r)為？

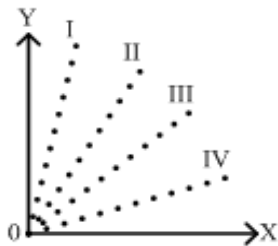
- (A) $r \neq 0.5$
- (B) $r = 0.5^*$

() 為什麼我會選擇上述的答案呢？下列哪一個選項**最**符合我的想法：

- ① 將兩變項對調，相關係數也會跟著改變
- ② 原本的相關係數是 $\frac{5}{10}$ ，將兩變數調換之後，相關係數就變成了 $\frac{10}{5}$
- ③ 運動時間可以直接影響健康指數，但是健康指數則會有其它的影響因素，因此健康指數與運動時間的相關係數會較小
- ④ 將兩變數對調，不會影響相關係數的值*
- ⑤ 其它_____

題 14 :

() 請你依據相關的強弱，將下圖中的 I 至 IV 從相關強度**最強者**排至相關強度**最弱者**。



- (A) $I > II > III > IV$
- (B) $(II = III) > (I = IV)$
- (C) $I = II = III = IV^*$

() 為什麼我會選擇上述的答案呢？下列哪一個選項**最**符合我的想法：

- ① 因為斜率為 1 是完全正相關，所以愈接近斜率為 1 的斜直線，相關強度愈強
- ② I, II, III, IV 都是一條斜直線，所以相關強度均相同*
- ③ 根據斜率大小判斷相關強度，斜率愈大者相關強度愈強
- ④ 其它_____

Bulletin of Educational Psychology, 2011, 42(3), 379-400
National Taiwan Normal University, Taipei, Taiwan, R.O.C.

Developing Two-Tier Diagnostic Instrument for Exploring Students' statistical misconceptions: Take "Correlation" as the Example

Tzu-Chien Liu

Center of Teacher Education
Graduate Institute of learning & Instruction
National Central University

Yi-Chun Lin

Graduate Institute of Learning & Instruction
National Central University

Statistical literacy is essential to people of modern era. However, research has shown that students often hold misconceptions when they study statistics. Although two-tier diagnostic instruments have been developed for diagnosing and exploring students' misconceptions in other fields of science, few are available in statistics education. Therefore, the main purpose of the current study was to develop a two tier diagnostic instrument: Two-tier Diagnostic Instrument about Correlation Misconceptions (TDICM). This instrument was adopted to analyze the responses of 982 12th-graders who had taken statistics courses to "correlation", and to explore the types, universality, and possible causes of such misconceptions in correlation. The research results showed that: (a) TDICM has satisfactory levels of difficulty, internal consistency reliability, and discrimination; (b) most participants do not have good comprehension of "correlation" concepts. Although some were able to answer the questions correctly in the first tier of TDICM, they failed to do so in the second tier due to lack of comprehensive understanding of correlation; (c) students' responses in the two-tier model indicated that there are 14 correlation misconceptions universally held by the participants; (d) seven of the 14 misconceptions has been discussed in the literature, and others were identified in this study. Furthermore, the possible causes of these misconceptions and the implications in education were discussed. Finally, this study provided some suggestions to the application of two-tier diagnostic instrument and the correlation misconceptions in statistics education.

KEY WORDS: correlation, misconception, statistics education, two-tier diagnostic instrument

